

SPECIFICATION

FLUORESCENT PROTEIN AND CHROMOPROTEIN

TECHNICAL FIELD

The present invention relates to a novel fluorescent protein. More specifically, the present invention relates to a novel fluorescent protein derived from *Montipora* sp., *Acropora* sp., and *Lobophytum crassum*, and the use thereof.

Further, the present invention relates to a novel chromoprotein. More specifically, the present invention relates to a novel chromoprotein derived from *Actinia equine*, and use thereof.

BACKGROUND ART

Green fluorescent protein (GFP) derived from *Aequorea victoria*, a jellyfish, has many purposes in biological systems. Recently, various GFP mutants have been produced based on the random mutagenesis and semi-rational mutagenesis, wherein a color is changed, a folding property is improved, luminance is enhanced, or pH sensitivity is modified. Fluorescent proteins such as GFP are fused with other proteins by gene recombinant technique, and monitoring of the expression and transportation of the fusion proteins is carried out.

One of the most commonly used types of GFP mutant is Yellow fluorescent protein (YFP). Among *Aequorea*-derived GFP mutants, YFP exhibits the fluorescence with the longest wavelength. The values ϵ and Φ of the majority of YFPs are 60,000 to 100,000 $M^{-1}cm^{-1}$ and 0.6 to 0.8, respectively (Tsien, R. Y. (1998). *Ann. Rev. Biochem.* 67, 509-544). These values are comparable to those of the general fluorescent group (fluorescein, rhodamine, etc.). Accordingly, improvement of the absolute luminance of YFP is nearly approaching its limit.

In addition, cyan fluorescent protein (CFP) is another example of the GFP mutant. Of this type of protein, ECFP (enhanced cyan fluorescent protein) has been known. Moreover, red fluorescent protein (RFP) has been isolated from sea anemone (*Discoma* sp.). Of this type of protein, DasRed has been known. Thus, 4 types of fluorescent proteins, that are, green fluorescent protein, yellow fluorescent protein, cyan fluorescent protein, and red fluorescent protein, have successively been developed. The range of the spectrum has significantly been expanded.

In addition, some cnidarians emit a fluorescence. The cloning of fluorescent protein genes derived from such cnidarians has been attempted. However, in order to expand the repertoire of fluorescent and biochemical properties, it is necessary to clone more genes.

On the other hand, a chromoprotein is obtained by bringing close to 0 (zero) the quantum yield of a conventional fluorescent protein. Such a chromoprotein can be applied to various purposes because it is able to introduce into a cell a molecule for converting light energy to other types of energy. However, there have been only a few reports regarding the absorption wavelength properties of such a chromoprotein.

DISCLOSURE OF THE INVENTION

It is an object of the present invention to provide a novel fluorescent protein derived from *Montipora* sp., *Acropora* sp., and *Lobophytum crassum*.

It is another object of the present invention to provide a fluorescent protein having a sharper spectrum than the wide excitation spectrum of the conventional RFP (DsRed; CLONTECH).

It is further another object of the present invention to provide a novel chromoprotein derived from *Actinia equina*, which absorbs a light with a specific wavelength.

The present inventors have conducted intensive studies directed towards

achieving the aforementioned objects. The inventors have designed preferred primers based on information regarding the amino acid sequences of known fluorescent proteins. Thereafter, using the aforementioned primers, the inventors have succeeded in amplifying genes encoding novel fluorescent proteins obtained from cDNA libraries derived from *Montipora* sp., *Acropora* sp., and *Lobophytum crassum*, and in cloning them. Moreover, the present inventors have analyzed the fluorescence properties and pH sensitivity of the obtained fluorescent proteins derived from *Montipora* sp., *Acropora* sp., and *Lobophytum crassum*. Furthermore, the inventors have designed preferred primers based on information regarding the amino acid sequences of known fluorescent proteins, and thereafter, using the aforementioned primers, the inventors have succeeded in amplifying a gene encoding a novel chromoprotein protein obtained from a cDNA library derived from *Actinia equina* presenting a red color, and in cloning it. Still further, the present inventors have analyzed the light absorption properties and pH sensitivity of the obtained chromoprotein derived from *Actinia equina*. The present invention has been completed based on these findings.

Thus, the present invention provides the the following (1) to (35):

(1) A fluorescent protein derived from *Montipora* sp., which has the following properties:

- [1] the excitation maximum wavelength is 507 nm;
- [2] the fluorescence maximum wavelength is 517 nm;
- [3] the molar absorption coefficient at 507 nm is 104,050;
- [4] the quantum yield is 0.29; and
- [5] the pH sensitivity of light absorption property is pKa of approximately 5.5.

(2) A fluorescent protein derived from *Acropora* sp., which has the following properties:

- [1] the excitation maximum wavelength is 505 nm;
- [2] the fluorescence maximum wavelength is 516 nm;

[3] the molar absorption coefficient at 505 nm is 53,600;

[4] the quantum yield is 0.67; and

[5] the pH sensitivity of light absorption property is pKa of approximately 6.4.

(3) A fluorescent protein derived from *Acropora* sp., which has the following properties:

[1] the excitation maximum wavelength is 472 nm;

[2] the fluorescence maximum wavelength is 496 nm;

[3] the molar absorption coefficient at 472 nm is 27,250;

[4] the quantum yield is 0.90; and

[5] the pH sensitivity of light absorption property is pKa of approximately 6.6.

(4) A fluorescent protein derived from *Montipora* sp., which has the following properties:

[1] the excitation maximum wavelength is 557 nm;

[2] the fluorescence maximum wavelength is 574 nm;

[3] the molar absorption coefficient at 557 nm is 41,750;

[4] the quantum yield is 0.41; and

[5] the pH sensitivity of light absorption property is pKa < approximately 4.0.

(5) A chromoprotein derived from *Actinia equina*, which has the following properties:

[1] the absorption maximum wavelength is 592 nm;

[2] the molar absorption coefficient at 592 nm is 87,000; and

[3] the pH sensitivity of light absorption property is stable in the range between pH 5 and 10.

(6) A fluorescent protein derived from *Lobophytum crassum*, which has the following properties:

[1] the excitation maximum wavelength is 482 nm;

[2] the fluorescence maximum wavelength is 498 nm;

[3] the molar absorption coefficient at 482 nm is 71,000;

[4] the quantum yield is 0.41; and

[5] the pH sensitivity of the fluorescence maximum is stable in the range between pH 4 and 10.

(7) A fluorescent protein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 1; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, and which has a fluorescence.

(8) A fluorescent protein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 3; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, and which has a fluorescence.

(9) A fluorescent protein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 5 or 7; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, and which has a fluorescence.

(10) A fluorescent protein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 9; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 9, and which has a fluorescence.

(11) A chromoprotein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 11; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of

one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 11, and which has light-absorbing properties.

(12) A fluorescent protein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 13; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 13, and which has fluorescence.

(13) DNA encoding the protein according to any one of (1) to (12) above.

(14) DNA of either the following (a) or (b):

(a) DNA encoding an amino acid sequence shown in SEQ ID NO: 1; or

(b) DNA which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, and which encodes a fluorescent protein.

(15) DNA having either the following nucleotide sequence (a) or (b):

(a) a nucleotide sequence shown in SEQ ID NO: 2; or

(b) a nucleotide sequence which comprises a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 2, and which encodes a fluorescent protein.

(16) DNA of either the following (a) or (b):

(a) DNA encoding an amino acid sequence shown in SEQ ID NO: 3; or

(b) DNA which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, and which encodes a fluorescent protein.

(17) DNA having either the following nucleotide sequence (a) or (b):

(a) a nucleotide sequence shown in SEQ ID NO: 4; or

(b) a nucleotide sequence which comprises a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO:

4, and which encodes a fluorescent protein.

(18) DNA of either the following (a) or (b):

(a) DNA encoding an amino acid sequence shown in SEQ ID NO: 5 or 7; or

(b) DNA which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, and which encodes a fluorescent protein.

(19) DNA having either the following nucleotide sequence (a) or (b):

(a) a nucleotide sequence shown in SEQ ID NO: 6 or 8; or

(b) a nucleotide sequence which comprises a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 6 or 8, and which encodes a fluorescent protein.

(20) DNA of either the following (a) or (b):

(a) DNA encoding an amino acid sequence shown in SEQ ID NO: 9; or

(b) DNA which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 9, and which encodes a fluorescent protein.

(21) DNA having either the following nucleotide sequence (a) or (b):

(a) a nucleotide sequence shown in SEQ ID NO: 10; or

(b) a nucleotide sequence which comprises a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 10, and which encodes a fluorescent protein.

(22) DNA of either the following (a) or (b):

(a) DNA encoding an amino acid sequence shown in SEQ ID NO: 11; or

(b) DNA which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 11, and which encodes a protein having light-absorbing properties.

(23) DNA having either the following nucleotide sequence (a) or (b):

- (a) a nucleotide sequence shown in SEQ ID NO: 12; or
 - (b) a nucleotide sequence which comprises a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 12, and which encodes a protein having light-absorbing properties.
- (24) DNA of either the following (a) or (b):
- (a) DNA encoding an amino acid sequence shown in SEQ ID NO: 13; or
 - (b) DNA which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 13, and which encodes a fluorescent protein.
- (25) DNA having either the following nucleotide sequence (a) or (b):
- (a) a nucleotide sequence shown in SEQ ID NO: 14; or
 - (b) a nucleotide sequence which comprises a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 14, and which encodes a fluorescent protein.
- (26) A recombinant vector having the DNA according to any one of (13) to (25) above.
- (27) A transformant having the DNA according to any one of (13) to (25) above or the recombinant vector according to (26) above.
- (28) A fusion fluorescent protein, which consists of the fluorescent protein according to any one of (1) to (4), (6), (7) to (10), and (12) above, and another protein.
- (29) The fusion fluorescent protein according to (28) above, wherein another protein is a protein that localizes in a cell.
- (30) The fusion fluorescent protein according to (28) or (29) above, wherein another protein is a protein specific to a cell organella.
- (31) A fusion protein, which consists of the chromoprotein according to (5) or (11) above and another protein.
- (32) A method for analyzing the localization or dynamics of a protein in a cell, which

is characterized in that the fusion fluorescent protein according to any one of (28) to (30) above is allowed to express in the cell.

(33) A method for analyzing physiologically active substances, which is characterized in that the FRET (fluorescence resonance energy transfer) method is carried out using the chromoprotein according to (5) or (11) above as an acceptor protein.

(34) A fluorescent reagent kit, which comprises: the fluorescent protein of any one of (1) to (4), (6), (7) to (10), and (12) above; the DNA of any one of (14) to (21), (24), and (25) above; the recombinant vector of (26) above; the transformant of (27) above; or the fusion fluorescent protein of any of (28) to (30).

(35) An absorbance reagent kit, which comprises: the chromoprotein of (5) or (11) above; the DNA of (22) or (23) above; the recombinant vector of (26) above; the transformant of (27) above; or the fusion protein of (31) above.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the fluorescence spectrum and excitation spectrum of the fluorescent protein (COG) of the present invention derived from *Montipora* sp. (Figure A), the absorption spectrum of the fluorescent protein (COG) (Figure B), and the pH sensitivity of the fluorescent protein (COG) (Figure C). In Figure C, the horizontal axis represents pH value, and the longitudinal axis represents absorbance.

Figure 2 shows the fluorescence spectrum and excitation spectrum at pH 5 of the fluorescent protein (COG) of the present invention derived from *Montipora* sp. (Figure A), and the absorption spectrum thereof at pH 5 (Figure B).

Figure 3 shows the fluorescence spectrum and excitation spectrum of the fluorescent protein (MIG) of the present invention derived from *Acropora* sp. (Figure A), the absorption spectrum of the fluorescent protein (MIG) (Figure B), and the pH sensitivity of the fluorescent protein (MIG) (Figure C). In Figure C, the horizontal axis represents pH value, and the longitudinal axis represents absorbance.

Figure 4 shows the fluorescence spectrum and excitation spectrum of the fluorescent protein (MICy) of the present invention derived from *Acropora* sp. (Figure A), the absorption spectrum of the fluorescent protein (MICy) (Figure B), and the pH sensitivity of the fluorescent protein (MICy) (Figure C). In Figure C, the horizontal axis represents pH value, and the longitudinal axis represents absorbance.

Figure 5 shows the pH sensitivity of the fluorescent protein (MiCy2) of the present invention (Figure A), and the excitation and fluorescence spectra thereof (Figure B).

Figure 6 shows the fluorescence spectrum and excitation spectrum of the fluorescent protein (COR) of the present invention derived from *Montipora* sp. (Figure A), the absorption spectrum of the fluorescent protein (COR) (Figure B), and the pH sensitivity of the fluorescent protein (COR) (Figure C). In Figure C, the horizontal axis represents pH value, and the longitudinal axis represents absorbance.

Figure 7 shows the results obtained by the measurement of the absorption spectrum (pH 7.9) of the chromoprotein (Ume) of the present invention derived from *Actinia equina* (Figure A), and the pH sensitivity of the absorption maximum of the chromoprotein (Ume) (Figure B). In Figure A, the horizontal axis represents the wavelength of an absorbed light, and the longitudinal axis represents absorbance. In Figure B, the horizontal axis represents pH value, and the longitudinal axis represents absorbance.

Figure 8 shows the fluorescence spectrum and excitation spectrum of the fluorescent protein (KnG) of the present invention derived from *Lobophytum crassum* (Figure A), and the pH dependence of the fluorescent protein (KnG) (Figure B).

BEST MODE FOR CARRYING OUT THE INVENTION

The embodiments of the present invention will be described in detail below.

(1) Fluorescent proteins and chromoprotein of the present invention

The first fluorescent protein of the present invention is derived from *Montipora* sp., and it has the following properties:

- [1] the excitation maximum wavelength is 507 nm;
- [2] the fluorescence maximum wavelength is 517 nm;
- [3] the molar absorption coefficient at 507 nm is 104,050;
- [4] the quantum yield is 0.29; and
- [5] the pH sensitivity of light absorption property is pKa of approximately 5.5.

Montipora sp. is one kind of coral, which belongs to Cnidaria, Anthozoa, Hexacorallia, Scleractinia, Artrocoeniidae. This coral often forms a massive or covered colony.

As described in examples below, the first fluorescent protein of the present invention has an excitation maximum wavelength of 507 nm and a fluorescence maximum wavelength of 517 nm. In addition, its molar absorption coefficient at 507 nm is 104,050, and its quantum yield is 0.29. The molar absorption coefficient represents the amount of photon absorbed per mole of a fluorescent molecule. The quantum yield is a numerical value representing the degree of the absorbed photon that can be emitted as a fluorescence.

A specific example of the first fluorescent protein of the present invention is a fluorescent protein having either the following amino acid sequence (a) or (b):

- (a) an amino acid sequence shown in SEQ ID NO: 1; or
- (b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, and which has a fluorescence.

The second fluorescent protein of the present invention is derived from *Acropora* sp., and it has the following properties:

- [1] the excitation maximum wavelength is 505 nm;
- [2] the fluorescence maximum wavelength is 516 nm;

[3] the molar absorption coefficient at 505 nm is 53,600;

[4] the quantum yield is 0.67; and

[5] the pH sensitivity of light absorption property is pKa of approximately 6.4.

Acropora sp. is one kind of coral, which belongs to Cnidaria, Anthozoa, Hexacorallia, Scleractinia, Artrocoeniidae. This coral often forms a ramiform or tabular colony.

As described in examples below, the second fluorescent protein of the present invention has an excitation maximum wavelength of 505 nm and a fluorescence maximum wavelength of 516 nm. In addition, its molar absorption coefficient at 505 nm is 53,600, and its quantum yield is 0.67. The molar absorption coefficient represents the amount of photon absorbed per mole of a fluorescent molecule. The quantum yield is a numerical value representing the degree of the absorbed photon that can be emitted as a fluorescence.

A specific example of the second fluorescent protein of the present invention is a fluorescent protein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 3; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 3, and which has a fluorescence.

The third fluorescent protein of the present invention is derived from Acropora sp., and it has the following properties:

[1] the excitation maximum wavelength is 472 nm;

[2] the fluorescence maximum wavelength is 496 nm;

[3] the molar absorption coefficient at 472 nm is 27,250;

[4] the quantum yield is 0.90; and

[5] the pH sensitivity of light absorption property is pKa of approximately 6.6.

As described in examples below, the third fluorescent protein of the present

invention has an excitation maximum wavelength of 472 nm and a fluorescence maximum wavelength of 496 nm. In addition, its molar absorption coefficient at 472 nm is 27,250, and its quantum yield is 0.90. The molar absorption coefficient represents the amount of photon absorbed per mole of a fluorescent molecule. The quantum yield is a numerical value representing the degree of the absorbed photon that can be emitted as a fluorescence.

A specific example of the third fluorescent protein of the present invention is a fluorescent protein having either the following amino acid sequence (a) or (b):

- (a) the amino acid sequence shown in SEQ ID NO: 5 or 7; or
- (b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 5 or 7, and which has fluorescence.

The fourth fluorescent protein of the present invention is derived from *Montipora* sp., and it has the following properties:

- [1] the excitation maximum wavelength is 557 nm;
- [2] the fluorescence maximum wavelength is 574 nm;
- [3] the molar absorption coefficient at 557 nm is 41,750;
- [4] the quantum yield is 0.41; and
- [5] the pH sensitivity of light absorption property is $pK_a < \text{approximately } 4.0$.

As described in examples below, the fourth fluorescent protein of the present invention has an excitation maximum wavelength of 557 nm and a fluorescence maximum wavelength of 574 nm. In addition, its molar absorption coefficient at 557 nm is 41,750, and its quantum yield is 0.41. The molar absorption coefficient represents the amount of photon absorbed per mole of a fluorescent molecule. The quantum yield is a numerical value representing the degree of the absorbed photon that can be emitted as a fluorescence.

A specific example of the fourth fluorescent protein of the present invention is a

fluorescent protein having either the following amino acid sequence (a) or (b):

(a) the amino acid sequence shown in SEQ ID NO: 9; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 9, and which has a fluorescence.

The chromoprotein of the present invention is derived from *Actinia equina*, and it has the following properties:

[1] the absorption maximum wavelength is 592 nm;

[2] the molar absorption coefficient at 592 nm is 87,000; and

[3] the pH sensitivity of light absorption property is stable in the range between pH 5 and 10.

Actinia equina is one kind of sea anemone, which belongs to Cnidaria, Anthozoa, Hexacorallia, Actiniaria, Actiniidae. This sea anemone is commonly found in surf zones in Kyushu and more northern areas in Japan. When *Actinia equina* extends its tentacles, it seems to be a red flow that opens in water.

In the examples of the present specification described below, the chromoprotein having the aforementioned properties has been isolated using *Actinia equina* as a starting material. However, there are some cases where the chromoprotein of the present invention can be obtained from sea anemone other than *Actinia equina*. Such chromoproteins are also included in the scope of the present invention.

As described in examples below, the chromoprotein of the present invention has an absorption maximum wavelength of 592 nm and a molar absorption coefficient of 87,000 at 592 nm.

The molar absorption coefficient represents the amount of photon absorbed per mole of a fluorescent molecule. The quantum yield is a numerical value representing the degree of the absorbed photon that can be emitted as a fluorescence. Since the quantum yield of the chromoprotein of the present invention is extremely low, it emits

almost no fluorescence. Because of such a property, the chromoprotein of the present invention can be used (1) as an acceptor molecule (energy receptor) in FRET; (2) for the development of a system for converting the energy of a light irradiated to energy other than the light energy; or (3) for introducing a mutation into the amino acid sequence of a protein, so as to modify it such that it emits a fluorescence.

In addition, the chromoprotein of the present invention is characterized in that its pH sensitivity of the light absorption property is stable in the range between pH 5 and 10. That is to say, in the case of the chromoprotein of the present invention, the peak value of its absorption spectrum does not fluctuate so much in the range between pH 5 and 10. Accordingly, the chromoprotein of the present invention can be used in a wide range of pH environment under the same above conditions, and it can be used in living bodies without rigid constraints.

A specific example of the chromoprotein of the present invention is a chromoprotein having either the following amino acid sequence (a) or (b):

(a) an amino acid sequence shown in SEQ ID NO: 11; or

(b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 11., and which has light-absorbing properties.

The fifth fluorescent protein of the present invention is derived from *Lobophytum crassum*, and it has the following properties:

[1] the excitation maximum wavelength is 482 nm;

[2] the fluorescence maximum wavelength is 498 nm;

[3] the molar absorption coefficient at 482 nm is 71,000;

[4] the quantum yield is 0.41; and

[5] the pH sensitivity of the fluorescence maximum is stable in the range between pH 4 and 10.

Lobophytum crassum is one kind of coral, which belongs to Cnidaria, Anthozoa,

Octocorallia.

As described in examples below, the fifth fluorescent protein of the present invention has an excitation maximum wavelength of 482 nm and a fluorescence maximum wavelength of 498 nm. In addition, its molar absorption coefficient at 482 nm is 71,000, and its quantum yield is 0.41. The molar absorption coefficient represents the amount of photon absorbed per mole of a fluorescent molecule. The quantum yield is a numerical value representing the degree of the absorbed photon that can be emitted as a fluorescence.

A specific example of the fifth fluorescent protein of the present invention is a fluorescent protein having either the following amino acid sequence (a) or (b):

- (a) an amino acid sequence shown in SEQ ID NO: 13; or
- (b) an amino acid sequence, which comprises a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 13., and which has a fluorescence.

The scope of "one or several" in the phrase "an amino acid sequence comprising a deletion, substitution and/or addition of one or several amino acids" used herein is not particularly limited. For example, it means 1 to 20, preferably 1 to 10, more preferably 1 to 7, further preferably 1 to 5, and particularly preferably 1 to 3.

The term "has a fluorescence" covers all of the cases where any fluorescence is given. Various properties such as fluorescence intensity, excitation wavelength, fluorescence wavelength or pH sensitivity, may be changed or may remain unchanged, as compared with the case of the protein having an amino acid sequence shown in SEQ ID NO: 1.

The term "light-absorbing properties" is used in the present specification to mean properties capable of absorbing a light with a certain wavelength. For example, as with the chromoprotein described in the present specification, an absorption maximum wavelength may be 592 nm, or the value of such an absorption maximum wavelength

may be shifted. The pH sensitivity of the light absorption property is preferably stable in the range between pH 5 and 10.

As stated above, the chromoprotein of the present invention having the amino acid sequence shown in SEQ ID NO: 11 of the sequence listing emits almost no fluorescence. In the present invention, a deletion, substitution, and/or addition of one or several amino acids may be introduced into the amino acid sequence shown in SEQ ID NO: 11, so as to produce a protein emitting a stronger fluorescence. The thus produced protein is also included in the scope of the present invention.

The method of obtaining the fluorescent protein and the chromoproteins of the present invention is not particularly limited. The proteins may be either a protein synthesized by chemosynthesis, or recombinant protein produced by a gene recombination technique.

Where a recombinant protein is produced, it is necessary to obtain DNA encoding the protein. Appropriate primers are designed by using information regarding the amino acid sequence shown in SEQ ID NO: 1, 3, 5, 7, 9, 11 or 13 of the sequence listing of the present specification and the nucleotide sequence shown in SEQ ID NO: 2, 4, 6, 8, 10, 12 or 14 thereof. Using these primers, PCR is carried out by using cDNA library derived from *Montipora* sp., *Acropora* sp., *Actinia equine*, or *Lobophytum crassum* as a template, so that DNA encoding the fluorescent protein or the chromoprotein of the present invention can be obtained. Where a partial fragment of DNA encoding the fluorescent protein or the chromoprotein of the present invention are obtained by the above-described PCR, the produced DNA fragments are ligated to one another by a gene recombination technique, so that DNA encoding the desired fluorescent protein or the chromoprotein can be obtained. The fluorescent protein or the chromoprotein of the present invention can be produced by introducing this DNA into an appropriate expression system. Expression in an expression system will be described later in the present specification.

(2) DNA of the present invention

The present invention provides genes encoding the fluorescent proteins or chromoprotein of the present invention.

A specific example of DNA encoding the fluorescent proteins of the present invention is either the following DNA (a) or (b):

(a) DNA encoding an amino acid sequence shown in SEQ ID NO: 1, 3, 5, 7, 9, or 13; or

(b) DNA, which encodes an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 1, 3, 5, 7, 9, or 13, and which encodes a fluorescent protein.

Another specific example of DNA encoding the fluorescent proteins of the present invention is either the following DNA (a) or (b):

(a) DNA having the nucleotide sequence shown in SEQ ID NO: 2, 4, 6, 8, 10, or 14; or

(b) DNA, which has a nucleotide sequence comprising a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO: 2, 4, 6, 8, 10, or 14, and which encodes a fluorescent protein.

A specific example of DNA encoding the chromoprotein of the present invention is either the following DNA (a) or (b):

(a) DNA encoding an amino acid sequence shown in SEQ ID NO: 11; or

(b) DNA, which has an amino acid sequence comprising a deletion, substitution, and/or addition of one or several amino acids with respect to the amino acid sequence shown in SEQ ID NO: 11, and which encodes a protein having light-absorbing properties.

Another specific example of DNA encoding the chromoprotein of the present invention is DNA having either the following nucleotide sequence (a) or (b):

(a) a nucleotide sequence shown in SEQ ID NO: 12; or

(b) a nucleotide sequence which comprises a deletion, substitution, and/or addition of one or several nucleotides with respect to the nucleotide sequence shown in SEQ ID NO:

12, and which encodes a protein having light-absorbing properties.

The DNA of the present invention can be synthesized by, for example, the phosphoramidite method, or it can also be produced by polymerase chain reaction (PCR) using specific primers. The DNA of the present invention or its fragment is produced by the method described above in the specification.

A method of introducing a desired mutation into a certain nucleic acid sequence is known to a person skilled in the art. For example, known techniques such as a site-directed mutagenesis, PCR using degenerated oligonucleotides, or the exposure of cells containing nucleic acid to mutagens or radioactive rays, are appropriately used, so as to construct DNA having a mutation. Such known techniques are described in, for example, *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY., 1989; and *Current Protocols in Molecular Biology*, Supplements 1 to 38, John Wiley & Sons (1987-1997).

(3) Recombinant vector of the present invention

The DNA of the present invention can be inserted into a suitable vector and used. The type of a vector used in the present invention is not particularly limited. For example, it may be either a vector that can autonomously replicate (e.g., a plasmid, etc.), or vector that is incorporated into the genomes of host cells when it is introduced into the host cells and is then replicated together with the chromosome into which it is incorporated.

The vector used in the present invention is preferably an expression vector. In an expression vector, elements necessary for transcription (e.g., a promoter, etc.) are functionally ligated to the DNA of the present invention. The promoter is a DNA sequence which shows a transcriptional activity in host cells, and it is appropriately selected depending on the type of host cells.

Examples of a promoter which can operate in bacterial cells may include a

Bacillus stearothermophilus maltogenic amylase gene promoter, a *Bacillus licheniformis* alpha-amylase gene promoter, a *Bacillus amyloliquefaciens* BAN amylase gene promoter, a *Bacillus subtilis* alkaline protease gene promoter, a *Bacillus pumilus* xylosidase gene promoter, P_R and P_L promoters of phage rhamda, and lac, trp and tac promoters of *Escherichia coli*.

Examples of a promoter which can operate in mammalian cells may include an SV40 promoter, an MT-1 (metallothionein gene) promoter, and an adenovirus-2 major late promoter. Examples of a promoter which can operate in insect cells may include a polyhedrin promoter, a P10 promoter, an *Autographa californica* polyhedrosis basic protein promoter, a baculovirus immediate-early gene 1 promoter, and a baculovirus 39K delayed-early gene promoter. Examples of a promoter which can be operate in yeast host cells may include promoters derived from yeast glycolytic genes, an alcohol dehydrogenase gene promoter, a TPI1 promoter, and an ADH2-4c promoter.

Examples of a promoter which can operate in filamentous cells may include an ADH3 promoter and a *tpiA* promoter.

In addition, an appropriate terminator such as a human growth hormone terminator, or a TPI1 terminator or ADH3 terminator for fungal cells, may be functionally bound to the DNA of the present invention, as necessary. The recombinant vector of the present invention may further have elements such as a polyadenylation signal (e.g., one derived from SV40 or the adenovirus 5E1b region), a transcription enhancer sequence (e.g., an SV40 enhancer), or a translation enhancer sequence (e.g., one encoding the adenovirus VA RNA).

The recombinant vector of the present invention may further comprise a DNA sequence which enables the replication of the recombinant vector in host cells. SV40 replication origin is an example of such a sequence (when the host cells are mammalian cells).

The recombinant vector of the present invention may further comprise a

selective marker. Examples of such a selective marker may include genes, complements of which are absent from host cells, such as a dihydrofolate reductase (DHFR) gene or a *Shizosaccharomyces pombe* TPI gene, and drug resistant genes such as ampicillin, kanamycin, tetracycline, chloramphenicol, neomycin or hygromycin-resistant genes.

A method for ligating the DNA of the present invention, a promoter and, as desired, a terminator and/or a secretory signal sequence to one another and inserting these items into a suitable vector is known to a person skilled in the art.

(4) Transformant of the present invention

A transformant can be produced by introducing the DNA or recombinant vector of the present invention into a suitable host.

Any cell can be used as a host cell into which the DNA or recombinant vector of the present invention is introduced, as long as the DNA construct of the present invention can be expressed therein. Examples of such a cell may include bacteria, yeasts, fungal cells, and higher eukaryotic cells.

Examples of bacteria may include Gram-positive bacteria such as *Bacillus* or *Streptomyces*, and Gram-negative bacteria such as *Escherichia coli*. These bacteria may be transformed by the protoplast method or other known methods, using competent cells.

Examples of mammalian cells may include HEK 293 cells, HeLa cells, COS cells, BHK cells, CHL cells, and CHO cells. A method of transforming mammalian cells and expressing the introduced DNA sequence in the cells is also known. Examples of such a method may include the electroporation, the calcium phosphate method, and the lipofection method.

Examples of yeast cells may include those belonging to *Saccharomyces* or *Shizosaccharomyces*. Examples of such cells may include *Saccharomyces cerevisiae*

and *Saccharomyces kluyveri*. Examples of a method of introducing a recombinant vector into yeast host cells may include the electroporation, the spheroplast method, and the lithium acetate method.

Examples of other fungal cells may include those belonging to *Filamentous fungi* such as *Aspergillus*, *Neurospora*, *Fusarium* or *Trichoderma*. Where *Filamentous fungi* are used as host cells, transformation can be carried out by incorporating DNA constructs into host chromosomes, so as to obtain recombinant host cells. Incorporation of DNA constructs into the host chromosomes is carried out by known methods, and such known methods may include homologous recombination and heterologous recombination.

Where insect cells are used as host cells, both a vector into which a recombinant gene is introduced and a baculovirus are co-introduced into insect cells, and a recombinant virus is obtained in the culture supernatant of the insect cells. Thereafter, insect cells are infected with the recombinant virus, so as to allow the cells to express proteins (described in, for example, *Baculovirus Expression Vectors, A Laboratory Manual*; and *Current Protocols in Molecular Biology, Bio/Technology*, 6, 47 (1988)).

The *Autographa californica* nuclear polyhedrosis virus, which is a virus infecting to insects belonging to *Barathra brassicae*, can be used as baculovirus.

Examples of insect cells used herein may include Sf9 and Sf21, which are *Spodoptera frugiperda* ovarian cells [*Baculovirus Expression Vectors, A Laboratory Manual*, W. H. Freeman & Company, New York, (1992)], and HiFive (manufactured by Invitrogen), which are *Trichoplusia ni* ovarian cells.

Examples of the method of co-introducing both a vector into which a recombinant gene has been introduced and the above baculovirus into insect cells to prepare a recombinant virus may include the calcium phosphate method and the lipofection method.

The above transformant is cultured in an appropriate nutritive medium under

conditions enabling the introduced DNA construct to be expressed. In order to isolate and purify the protein of the present invention from the culture product of the transformant, common methods of isolating and purifying proteins may be used.

For example, where the protein of the present invention is expressed in a state dissolved in cells, after completion of the culture, cells are recovered by centrifugal separation, and the recovered cells are suspended in a water type buffer. Thereafter, the cells are disintegrated using an ultrasonic disintegrator or the like, so as to obtain a cell-free extract. A supernatant is obtained by centrifuging the cell-free extract, and then, a purified sample can be obtained from the supernatant by applying, singly or in combination, the following ordinary protein isolation and purification methods: the solvent extraction, the salting-out method using ammonium sulfate or the like, the desalting method, the precipitation method using an organic solvent, the anion exchange chromatography using resins such as diethylaminoethyl (DEAE) sepharose, the cation exchange chromatography using resins such as S-Sepharose FF (manufactured by Pharmacia), the hydrophobic chromatography using resins such as butyl sepharose or phenyl sepharose, the gel filtration method using a molecular sieve, the affinity chromatography, the chromatofocusing method, and the electrophoresis such as isoelectric focusing.

(5) Use of the fluorescent protein of the present invention and a fusion fluorescent protein comprising the same

The fluorescent protein of the present invention can be fused with another protein, so as to construct a fusion fluorescent protein.

A method of obtaining the fusion fluorescent protein of the present invention is not particularly limited. It may be either a protein synthesized by chemosynthesis, or recombinant protein produced by a gene recombination technique.

Where a recombinant protein is produced, it is necessary to obtain DNA

encoding the protein. Appropriate primers are designed using the information regarding the amino acid sequence shown in SEQ ID NO: 1, 3, 5, 7, 9 or 13 of the sequence listing of the present specification and the nucleotide sequence shown in SEQ ID NO: 2, 4, 6, 8, 10 or 14 thereof. Using these primers, PCR is carried out using a DNA fragment containing the gene of the fluorescent protein of the present invention as a template, so as to produce DNA fragments necessary for construction of the DNA encoding the fluorescent protein of the present invention. Moreover, DNA fragment encoding a protein to be fused is also obtained in the same above manner.

Subsequently, the thus obtained DNA fragments are ligated to one another by a gene recombination technique, so that DNA encoding the desired fusion fluorescent protein can be obtained. This DNA is then introduced into an appropriate expression system, so that the fusion fluorescent protein of the present invention can be produced.

The fluorescent protein of the present invention has an extremely high utility value as a marker. This is to say, the fluorescent protein of the present invention is purified as a fusion protein with an amino acid sequence to be tested, and the fusion protein is introduced into cells by methods such as the microinjection. By observing the distribution of the fusion protein over time, targeting activity of the amino acid sequence to be tested can be detected in the cells.

The type of another protein (an amino acid sequence to be tested) with which the fluorescent protein of the present invention is fused is not particularly limited. Preferred examples may include proteins localizing in cells, proteins specific for intracellular organelles, and targeting signals (e.g., a nuclear transport signal, a mitochondrial presequence, etc.). In addition, the fluorescent protein of the present invention can be expressed in cells and used, as well as being introduced into cells by the microinjection or the like. In this case, a vector into which the DNA encoding the fluorescent protein of the present invention is inserted in such a way that it can be expressed, is introduced into host cells.

Moreover, the fluorescent protein of the present invention can also be used as a reporter protein to determine promoter activity. This is to say, a vector is constructed such that DNA encoding the fluorescent protein of the present invention is located downstream of a promoter to be tested, and the vector is then introduced into host cells. By detecting the fluorescence of the fluorescent protein of the present invention which is emitted from the cells, the activity of the promoter to be tested can be determined. The type of a promoter to be tested is not particularly limited, as long as it operates in host cells.

A vector used to detect the targeting activity of the above amino acid sequence to be tested or to determine promoter activity is not particularly limited. Examples of a vector preferably used for animal cells may include pNEO (P. Southern, and P. Berg (1982) J. Mol. Appl. Genet. 1: 327), pCAGGS (H. Niwa, K. Yamamura, and J. Miyazaki, Gene 108, 193-200 (1991)), pRc/CMV (manufactured by Invitrogen), and pCDM8 (manufactured by Invitrogen). Examples of a vector preferably used for yeasts may include pRS303, pRS304, pRS305, pRS306, pRS313, pRS314, pRS315, pRS316 (R. S. Sikorski and P. Hieter (1989) Genetics 122: 19-27), pRS423, pRS424, pRS425, pRS426 (T. W. Christianson, R. S. Sikorski, M. Dante, J. H. Shero, and P. Hieter (1992) Gene 110: 119-122).

In addition, the type of cells used herein is also not particularly limited. Various types of animal cells such as L cells, BalbC-3T3 cells, NIH3T3 cells, CHO (Chinese hamster ovary) cells, HeLa cells or NRK (normal rat kidney) cells, yeast cells such as *Saccharomyces cerevisiae*, *Escherichia coli* cells, or the like can be used. Vector can be introduced into host cells by common methods such as the calcium phosphate method or the electroporation.

The above obtained fusion fluorescent protein of the present invention wherein the fluorescent protein of the present invention is fused with another protein (referred to as a protein X) is allowed to be expressed in cells. By monitoring a fluorescence

emitted, it becomes possible to analyze the localization or dynamics of the protein X in cells. That is, cells transformed or transfected with DNA encoding the fusion fluorescent protein of the present invention are observed with a fluorescence microscope, so that the localization and dynamics of the protein X in the cells can be visualized and thus analyzed.

For example, by using a protein specific for an intracellular organella as a protein X, the distribution and movement of a nucleus, a mitochondria, an endoplasmic reticulum, a Golgi body, a secretory vesicle, a peroxisome, etc., can be observed.

Moreover, for example, axis cylinders or dendrites of the nerve cells show an extremely complicated change in strikes in an individual who is under development. Accordingly, fluorescent labeling of these sites enables a dynamic analysis.

The fluorescence of the fluorescent protein of the present invention can be detected with a viable cell. Such detection can be carried out using, for example, a fluorescence microscope (Axiophoto Filter Set 09 manufactured by Carl Zeiss) or an image analyzer (Digital Image Analyzer manufactured by ATTO).

The type of a microscope can be appropriately selected depending on purposes. Where frequent observation such as pursuit of a change over time is carried out, an ordinary incident-light fluorescence microscope is preferable. Where observation is carried out while resolution is emphasized, for example, in the case of searching localization in cells specifically, a confocal laser scanning microscope is preferable. In terms of maintenance of the physiological state of cells and prevention from contamination, an inverted microscope is preferable as a microscope system. When an erecting microscope with a high-powered lens is used, a water immersion lens can be used.

A filter set can be appropriately selected depending on the fluorescence wavelength of a fluorescent protein. In the case of the first and second fluorescent proteins of the present invention, a filter having an excitation light between

approximately 490 and 510 nm and a fluorescence between approximately 510 and 530 nm is preferably used. In the case of the third fluorescent proteins of the present invention, a filter having an excitation light between approximately 460 and 480 nm and a fluorescence between approximately 480 and 510 nm is preferably used. In the case of the fourth fluorescent protein of the present invention, a filter having an excitation light between approximately 550 and 565 nm and a fluorescence between approximately 570 and 580 nm is preferably used. In the case of the fifth fluorescent protein of the present invention, a filter having an excitation light between approximately 470 and 490 nm and a fluorescence between approximately 490 and 510 nm is preferably used.

When viable cells are observed over time using a fluorescence microscope, a high sensitive cooled CCD camera is used, since photography is carried out in a short time. In the case of the cooled CCD camera, CCD is cooled to decrease thermal noise, so that a weak fluorescence image can be clearly photographed by exposure in a short time.

(6) Use of the chromoprotein of the present invention and a fusion protein comprising the same

The chromoprotein of the present invention can be fused with another protein, so as to construct a fusion protein. The type of said another protein to be fused to the chromoprotein of the present invention is not particularly limited, and preferred examples may include a protein which interacts with another molecule. The examples may include a receptor protein or ligand thereof, antigen, antibody and the like.

A method of obtaining the fusion protein of the present invention is not particularly limited. It may be either a protein synthesized by chemosynthesis, or recombinant protein produced by a gene recombination technique.

Where a recombinant fusion protein is produced, it is necessary to obtain DNA encoding the protein. The DNA encoding the chromoprotein of the present invention

and the DNA encoding the another protein to be fused to the chromoprotein, can be obtained by the method as mentioned above in this specification or by the method similar to it. Then, these DNA fragments are ligated to one another by a gene recombination technique, so that DNA encoding the desired fusion protein can be obtained. This DNA is then introduced into an appropriate expression system, so that the fusion protein of the present invention can be produced.

FRET (fluorescence resonance energy transfer) has been known as a means for analyzing the interaction between molecules. In FRET, for example, a first molecule labeled with a cyan fluorescent protein (CFP) acting as a first fluorescent protein is allowed to coexist with a second molecule labeled with a yellow fluorescent protein (YFP) acting as a second fluorescent protein, so as to allow the yellow fluorescent protein (YFP) to act as an acceptor molecule and to allow the cyan fluorescent protein (CFP) to act as a donor molecule. Thus, FRET (fluorescence resonance energy transfer) is allowed to take place between both molecules, so as to visualize the interaction between the first and second molecules. Namely, in FRET, different dyes are introduced into two types of molecules. One dyes with a higher energy level (a donor molecule) is selectively excited, and the fluorescence of the dye is measured. Long-wavelength fluorescence from the other dye (an acceptor molecule) is also measured. The interaction between the molecules is visualized by using the difference between the amounts of both fluorescences. Only when both dyes are adjacent to each other due to the interaction of the two types of molecules, a decrease in the fluorescence of the donor molecule and an increase in the fluorescence of the acceptor molecule are observed by single wavelength excitation dual wavelength photometry. However, in a case where a chromoprotein is used as an acceptor molecule, a decrease in the fluorescence of the donor molecule occurs only when both dyes are adjacent to each other by the interaction of the two types of molecules. Such a decrease can be observed

by single wavelength excitation single wavelength photometry. Thus, the use of the chromoprotein of the present invention enables facilitation of measurement apparatuses.

The chromoprotein of the present invention is particularly advantageous when it is used as an acceptor molecule in FRET (fluorescence resonance energy transfer). That is to say, a fused form (a first fused form) of the chromoprotein of the present invention and a test substance is first produced. Then, a fused form (a second fused form) of another test substance interacting with the above test substance and another fluorescent protein is produced. Thereafter, the first fused form is allowed to interact with the second fused form, and the generated fluorescence is analyzed, so that the interaction between the aforementioned two types of test substances can be analyzed. FRET (fluorescence resonance energy transfer) using the chromoprotein of the present invention may be carried out either in a test tube or in a cell.

(7) Kit of the present invention

The present invention provides a kit for analyzing the localization of intracellular components and/or analyzing physiologically active substances, which is characterized in that it comprises at least one selected from the fluorescent protein, the fusion fluorescent protein, the DNA, the recombinant vector, or the transformant, which are described in the present specification. The kit of the present invention can be produced from commonly used materials that are known per se, by using common methods.

Further, the present invention provides a light-absorbing reagent kit, which is characterized in that it comprises at least one selected from the chromoprotein, fusion protein, DNA, recombinant vector or transformant, which are described in the present specification. The kit of the present invention can be produced from commonly used materials that are known per se, by using common methods.

Reagents such as the fluorescent protein, the chromoprotein or the DNA are

dissolved in an appropriate solvent, so that the reagents can be prepared in a form suitable for conservation. Water, ethanol, various types of buffer solution, etc. can be used as such a solvent.

The present invention will be further described in the following examples. However, the present invention is not limited by these examples.

EXAMPLES

Example 1: Isolation of novel fluorescent protein gene (COG) from stony coral, and analysis of fluorescence properties thereof

(1) Extraction of total RNA

A fluorescent protein gene was isolated from coral. *Montipora* sp. was used as a material. A frozen *Montipora* sp. was crushed in a mortar, and 7.5 ml of "TRIzol" (GIBCO BRL) was then added to 2 g (wet weight) of the crushed *Montipora* sp. Thereafter, the obtained mixture was homogenized and then centrifuged at 1,500 x g for 10 minutes. Thereafter, 1.5 ml of chloroform was added to the obtained supernatant, and the mixture was then stirred for 15 seconds. Thereafter, the mixture was left at rest for 3 minutes. The resultant was then centrifuged at 7,500 x g for 15 minutes. Thereafter, 3.75 ml of isopropanol was added to the obtained supernatant, and the mixture was then stirred for 15 seconds, followed by leaving the mixture at rest for 10 minutes. Thereafter, the resultant was centrifuged at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and 6 ml of 70% ethanol was added to the residue, followed by centrifugation at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and the precipitate was then dissolved in 200 µl of DEPC water. Total RNA dissolved in the DEPC water was 100 times diluted, and the values of O.D.260 and O.D.280 were then measured, so as to determine RNA concentration. As a result, 22 µg of total RNA was obtained.

(2) Synthesis of first strand cDNA

cDNA (33 μ l) was synthesized from 4 μ g of the total RNA, using a kit for synthesizing first strand cDNA "Ready To Go" (Amersham Pharmacia).

(3) Degenerated PCR

3 μ l of the synthesized first strand cDNA (33 μ l) was used as a template to carry out PCR.

Primers were produced by making comparison among the amino acid sequences of known fluorescent proteins, extracting similar portions, and converting them to nucleotide sequences.

Primers used:

5'-GAAGGRTGYGTCAAYGGRCAY-3' (primer 1) (SEQ ID NO: 15)

5'-ACVGGDCCATYDGVAAGAAARTT-3' (primer 2) (SEQ ID NO: 16)

I represents inosine; R represents A or G; Y represents C or T; V represents A, C, or G; D represents A, G, or T; S represents C or G; H represents A, T, or C

Composition of PCR reaction solution:

| | |
|-------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 1 | 1 μ l |
| 100 μ M primer 2 | 1 μ l |
| Milli-Q | 35 μ l |
| taq polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 min (PAD)

94°C x 30 sec (denaturation)

52°C x 30 sec (annealing of primers to template)

72°C x 1 min (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 35 times.

72°C x 7 min (final elongation)

4°C (maintenance)

Using 1 µl of the amplified product obtained in the first PCR reaction as a template, PCR was carried out again under the same above conditions. A 350-bp fragment was cut out via agarose gel electrophoresis, and it was then purified.

(4) Subcloning and sequencing

The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer. Thereafter, the obtained nucleotide sequence was compared with the nucleotide sequences of other fluorescent protein genes, so as to determine whether the nucleotide sequence of the DNA fragment was derived from a fluorescent protein. With regard to those that were determined to be a part of the fluorescent protein genes, the full-length genes were cloned by the 5'-RACE method and the 3'-RACE method.

(5) 5'-RACE method

In order to determine the nucleotide sequence on the 5'-side of the DNA fragment obtained by the degenerated PCR, the 5'-RACE method was applied using 5'-RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (GIBCO BRL). 5 µg of the total RNA prepared in (1) above was used as a template.

The following primers were used in the first amplification of dC-tailed cDNA:

5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3' (primer 3) (SEQ ID NO: 17); and

5'-CCATCTTCAAAGAGAAAAGACCTTT-3' (primer 4) (SEQ ID NO: 18)

Herein, I represents inosine.

The following primers were used in the second amplification:

5'-GGCCACGCGTCGACTAGTAC-3' (primer 5) (SEQ ID NO: 19); and

5'-CATGAGTTCTTGAAATAGTCAAC-3' (primer 6) (SEQ ID NO: 20). PCR reaction conditions were applied in accordance with protocols attached to the kit.

The amplified 350-bp band was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

(6) 3'-RACE method

The nucleotide sequence on the 3'-side of the DNA fragment obtained by the degenerated PCR was obtained by PCR using a primer prepared based on the information obtained by determination of the nucleotide sequence in (4) above and an oligo dT primer. 3 μ l of the first strand cDNA prepared in (2) above was used as a template. The prepared primer was 5'-ATGGCTCTTTCAAAGCGAGGTG-3' (primer 7) (SEQ ID NO: 21).

Composition of PCR reaction solution:

| | |
|-------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 20 μ M primer 7 | 1 μ l |
| 10 μ M oligo dT primer | 1 μ l |
| Milli-Q | 35 μ l |
| Taq polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of approximately 1,000 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

The obtained full-length nucleotide sequence is shown in SEQ ID NO: 2 of the sequence listing, and the obtained full-length amino acid sequence is shown in SEQ ID NO: 1 thereof. This clone was named COG.

(7) Expression of protein in *Escherichia coli*

Based on the obtained full-length nucleotide sequence, a primer was produced with a portion corresponding to the N-terminus of the protein. An oligo dT primer was used as a primer corresponding to the C-terminal side thereof. Thereafter, using such primers, PCR was carried out employing the first strand cDNA prepared in (2) above as a template.

Primer used:

5'-GGGGGATCCGACCATGGCTCTTTCAAAGCGAGGTG-3' (primer 8) (SEQ ID NO: 22)

Composition of PCR reaction solution:

| | |
|------------------------------|------|
| Template (first strand cDNA) | 3 µl |
| X10 pyrobest buffer | 5 µl |
| 2.5 mM dNTPs | 4 µl |
| 100 µM primer 8 | 1 µl |

100 μ M oligo dT primer 1 μ l
Milli-Q 35 μ l
Pyrobest polymerase (5 U/ μ l) 1 μ l

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of approximately 1,000 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was subcloned into the BamHI-EcoRI site of a pRSET vector (Invitrogen), and it was then allowed to express in *Escherichia coli* (JM109-DE3). The expressed protein was constructed such that His-tag was attached to the N-terminus thereof, and thus it was purified with Ni-Agarose gel (QIAGEN). Purification was carried out in accordance with the attached protocols. Subsequently, the properties of the purified protein were analyzed.

(8) Analysis of fluorescence properties

Using a solution comprising 20 μ M fluorescent protein (COG), 150 mM KCl, and 50 mM HEPES(pH 7.5), the absorption spectrum of the protein was measured (Figure 1B). Thereafter, the molar absorption coefficient of the protein was calculated from the value of the peak (507 nm) of the spectrum. The fluorescent protein was diluted with the above buffer solution such that the absorption at 450 nm became 0.002. Its fluorescence spectrum by exciting at 450 nm and its excitation spectrum by a fluorescence at 550 nm were measured (Figure 1A). Likewise, EGFP (CLONTECH)

was diluted such that the absorption at 450 nm became 0.002, and its fluorescence spectrum was measured. Setting the quantum yield of EGFP to 0.6, the quantum yield of the cloned fluorescent protein was obtained. The results are shown in Table 1.

Table 1

| | Excitation maximum | Fluorescence maximum | Molar absorption coefficient | Quantum yield | pH sensitivity | Number of amino acids |
|-----|--------------------|----------------------|------------------------------|---------------|----------------|-----------------------|
| COG | 507 nm | 517 nm | 104,050 (507 nm) | 0.29 | pKa = 5.5 | 227 |

(9) Measurement of pH sensitivity

The fluorescent protein was diluted with each of the following buffer solutions to the same concentration. The value of the absorption at 570 nm was determined, and thus the pH sensitivity thereof was measured (Figure C1). The pH levels of the buffer solutions are as follows.

pH 4 and 5: Acetate buffer

pH 6 and 11: Phosphate buffer

pH 7 and 8: HEPES buffer

pH 9 and 10: Glycine buffer

This fluorescent protein has such characteristics that when pH 5 is compared with pH 6 to 10, the absorption peak is shifted to a short wavelength side (from 507 nm to 493 nm), and the fluorescence peak is also shifted to a short wavelength side (from 517 nm to 508 nm). The measurement results are shown in Figures 2A and 2B.

Example 2: Isolation of novel fluorescent protein gene (MIG) from coral

(1) Extraction of total RNA

A fluorescent protein gene was isolated from coral emitting a fluorescence.

Acropora sp. was used as a material. Acropora sp. was crushed with a hammer, and 15 ml of "TRIzol" (GIBCO BRL) was then added to 5 g of the crushed Montipora sp. Thereafter, the obtained mixture was stirred and then centrifuged at 1,500 x g for 10 minutes. Thereafter, 3 ml of chloroform was added to the obtained supernatant, and the mixture was then stirred for 15 seconds. Thereafter, the mixture was left at rest for 3 minutes. The resultant was then centrifuged at 7,500 x g for 15 minutes. Thereafter, 7.5 ml of isopropanol was added to the obtained supernatant, and the mixture was then stirred for 15 seconds, followed by leaving the mixture at rest for 10 minutes. Thereafter, the resultant was centrifuged at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and 6 ml of 70% ethanol was added to the residue, followed by centrifugation at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and the precipitate was then dissolved in 200 µl of DEPC water. Total RNA dissolved in the DEPC water was 100 times diluted, and the values of O.D.260 and O.D.280 were then measured, so as to determine RNA concentration. As a result, 220 µg of total RNA was obtained.

(2) Synthesis of first strand cDNA

cDNA (33 µl) was synthesized from 5 µg of the total RNA, using a kit for synthesizing first strand cDNA "Ready To Go" (Amersham Pharmacia).

(3) Degenerated PCR

3 µl of the synthesized first strand cDNA (33 µl) was used as a template to carry out PCR. Primers were produced by making comparison among the amino acid sequences of known fluorescent proteins, extracting similar portions, and converting them to nucleotide sequences.

Primers used:

5'-GAAGGRTGYGTCAAYGGRCAY-3' (primer 1) (SEQ ID NO: 15)

5'-ACVGGDCCATYDGVAAGAAARTT-3' (primer 2) (SEQ ID NO: 16)

R represents A or G; Y represents C or T; V represents A, C, or G; D represents A, G, or T

Composition of PCR reaction solution:

| | |
|-------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 1 | 1 μ l |
| 100 μ M primer 2 | 1 μ l |
| Milli-Q | 35 μ l |
| taq polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 min (PAD)

94°C x 30 sec (denaturation)

52°C x 30 sec (annealing of primers to template)

72°C x 1 min (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times. The annealing temperature was decreased by 0.3°C for every cycle. The annealing temperature was 43°C at the time of the 30th cycle.

72°C x 7 min (final elongation)

4°C (maintenance)

Using 1 μ l of the amplified product obtained in the first PCR reaction as a template, PCR was carried out again under the same above conditions. A band with an estimated size of 350 bp was cut out via agarose gel electrophoresis, and it was then purified.

(4) Subcloning and sequencing

The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA

sequencer. Thereafter, the obtained nucleotide sequence was compared with the nucleotide sequences of other fluorescent protein genes, so as to determine whether the nucleotide sequence of the DNA fragment was derived from a fluorescent protein. With regard to those that were determined to be a part of the fluorescent protein genes, the full-length genes were cloned by the 5'-RACE method and the 3'-RACE method.

(5) 5'-RACE method

In order to determine the nucleotide sequence on the 5'-side of the DNA fragment obtained by the degenerated PCR, the 5'-RACE method was applied using 5'-RAC System for Rapid Amplification of cDNA Ends, Version 2.0 (GIBCO BRL). 3 µg of the total RNA prepared in (1) above was used as a template. The following primers were used in the first amplification of dC-tailed cDNA:

5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3' (primer 3) (SEQ ID NO: 17); and

5'-TAGAAATGACCTTTCATATGACATTC-3' (primer 4) (SEQ ID NO: 23).

Herein, I represents inosine.

The following primers were used in the second amplification:

5'-GGCCACGCGTCGACTAGTAC-3' (primer 5) (SEQ ID NO: 19); and

5'-TCTGTTTCCATATTGAAAGGCTG-3' (primer 6) (SEQ ID NO: 24). PCR reaction conditions were applied in accordance with protocols attached to the kit.

The amplified 500-bp band was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

(6) 3'-RACE method

The nucleotide sequence on the 3'-side of the DNA fragment obtained by the

degenerated PCR was obtained by PCR, using a primer prepared based on the information obtained by determination of the nucleotide sequence in (4) above and an oligo dT primer. 3 µl of the first strand cDNA prepared in (2) above was used as a template. The prepared primer was

5'-ATGGTGTCTTATTCAAAGCAAGGCATCGCACA-3' (primer 7) (SEQ ID NO: 25).

Composition of PCR reaction solution:

| | |
|------------------------------|-------|
| Template (first strand cDNA) | 3 µl |
| X10 taq buffer | 5 µl |
| 2.5 mM dNTPs | 4 µl |
| 20 µM primer 7 | 1 µl |
| 10 µM oligo dT primer | 1 µl |
| Milli-Q | 35 µl |
| Taq polymerase (5 U/µl) | 1 µl |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

55°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of 900 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

The obtained full-length nucleotide sequence is shown in SEQ ID NO: 4 of the sequence listing, and the obtained full-length amino acid sequence is shown in SEQ ID NO: 3 thereof. This clone was named MIG.

(7) Expression of protein in Escherichia coli

Using a primer produced with a portion corresponding to the N-terminus of the obtained full-length nucleotide sequence of the protein and an oligo dT primer, PCR was carried out employing the first strand cDNA prepared in (2) above as a template.

Primer used:

5'-CGGGATCCGACCATGGTGTCTTATTCAAAGCAAGGCATCGCACA-3'(primer 8)
(SEQ ID NO: 26)

Composition of PCR reaction solution:

| | |
|------------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 pyrobest buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 20 μ M primer 8 | 1 μ l |
| 20 μ M oligo dT primer | 1 μ l |
| Milli-Q | 35 μ l |
| Pyrobest polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

55°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of 900 bp was cut out of the gel via agarose

gel electrophoresis and then purified. The purified DNA fragment was subcloned into the BamHI-EcoRI site of a pRSET vector (Invitrogen), and it was then allowed to express in *Escherichia coli* (JM109-DE3). The expressed protein was constructed such that His-tag was attached to the N-terminus thereof, and thus it was purified with Ni-Agarose gel (QIAGEN). Purification was carried out in accordance with the attached protocols. Subsequently, the properties of the purified protein were analyzed.

(8) Analysis of fluorescence properties

Using a solution comprising 20 μ M fluorescent protein (MIG), 50 mM KCl, and 50 mM HEPES(pH 7.4), the absorption spectrum of the protein was measured (Figure 3B). Thereafter, the molar absorption coefficient of the protein was calculated from the value of the peak (505 nm) of the spectrum. The fluorescent protein was diluted with the above buffer solution such that the absorption at 440 nm became 0.001. Its fluorescence spectrum by exciting at 440 nm and its excitation spectrum by a fluorescence at 540 nm were measured (Figure 3A). Likewise, EGFP (CLONTECH) was diluted such that the absorption at 440 nm became 0.001, and its fluorescence spectrum was measured. Setting the quantum yield of EGFP to 0.6, the quantum yield of the cloned fluorescent protein was obtained. The measurement results are shown in Table 2.

Table 2

| | Excitation maximum | Fluorescence maximum | Molar absorption coefficient | Quantum yield | pH sensitivity | Number of amino acids |
|-----|-----------------------|-------------------------|------------------------------------|------------------|-------------------|--------------------------|
| MIG | 505 nm | 516 nm | 53,600 (505 nm) | 0.67 | pKa = 6.4 | 232 |

(9) Measurement of pH sensitivity

The fluorescent protein was diluted with each of the following buffer solutions.

The value of the absorption at 505 nm was determined, and thus the pH sensitivity thereof was measured. The pH levels of the buffer solutions are as follows.

pH 4 and 5: Acetate buffer

pH 6 and 11: Phosphate buffer

pH 7 and 8: HEPES buffer

pH 9 and 10: Glycine buffer

The measurement results are shown in Figure 3C.

Example 3: Isolation of novel fluorescent protein gene (MICy) from coral

(1) Extraction of total RNA

A fluorescent protein gene was isolated from coral emitting a fluorescence. *Acropora* sp. was used as a material. *Acropora* sp. was crushed with a hammer, and 15 ml of "TRIzol" (GIBCO BRL) was then added to 5 g of the crushed *Acropora* sp. Thereafter, the obtained mixture was stirred and then centrifuged at 1,500 x g for 10 minutes. Thereafter, 3 ml of chloroform was added to the obtained supernatant, and the mixture was then stirred for 15 seconds. Thereafter, the mixture was left at rest for 3 minutes. The resultant was then centrifuged at 7,500 x g for 15 minutes. Thereafter, 7.5 ml of isopropanol was added to the obtained supernatant, and the mixture was then stirred for 15 seconds, followed by leaving the mixture at rest for 10 minutes. Thereafter, the resultant was centrifuged at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and 6 ml of 70% ethanol was added to the residue, followed by centrifugation at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and the precipitate was then dissolved in 200 µl of DEPC water. Total RNA dissolved in the DEPC water was 100 times diluted, and the values of O.D.260 and O.D.280 were then measured, so as to determine RNA concentration. As a result, 220 µg of total RNA was obtained.

(2) Synthesis of first strand cDNA

cDNA (33 μ l) was synthesized from 5 μ g of the total RNA, using a kit for synthesizing first strand cDNA “Ready To Go” (Amersham Pharmacia).

(3) Degenerated PCR

3 μ l of the synthesized first strand cDNA (33 μ l) was used as a template to carry out PCR. Primers were produced by making comparison among the amino acid sequences of known fluorescent proteins, extracting similar portions, and converting them to nucleotide sequences.

Primers used:

5'-GAAGGRTGYGTCAAYGGRCAY-3' (primer 1) (SEQ ID NO: 15)

5'-ACVGGDCCATYDGVAAAGAAARTT-3' (primer 2) (SEQ ID NO: 16)

R represents A or G; Y represents C or T; V represents A, C, or G; D represents A, G, or T

Composition of PCR reaction solution:

| | |
|-------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 1 | 1 μ l |
| 100 μ M primer 2 | 1 μ l |
| Milli-Q | 35 μ l |
| taq polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 min (PAD)

94°C x 30 sec (denaturation)

52°C x 30 sec (annealing of primers to template)

72°C x 1 min (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times. The annealing temperature was decreased by 0.3°C for every cycle. The annealing temperature was 43°C at the time of the 30th cycle.

72°C x 7 min (final elongation)

4°C (maintenance)

Using 1 µl of the amplified product obtained in the first PCR reaction as a template, PCR was carried out again under the same above conditions. A band with an estimated size of 350 bp was cut out via agarose gel electrophoresis, and it was then purified.

(4) Subcloning and sequencing

The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer. Thereafter, the obtained nucleotide sequence was compared with the nucleotide sequences of other fluorescent protein genes, so as to determine whether the nucleotide sequence of the DNA fragment was derived from a fluorescent protein. With regard to those that were determined to be a part of the fluorescent protein genes, the full-length genes were cloned by the 5'-RACE method and the 3'-RACE method.

(5) 5'-RACE method

In order to determine the nucleotide sequence on the 5'-side of the DNA fragment obtained by the degenerated PCR, the 5'-RACE method was applied using 5'-RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (GIBCO BRL). 3 µg of the total RNA prepared in (1) above was used as a template. The following primers were used in the first amplification of dC-tailed cDNA:

5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3' (primer 3) (SEQ ID NO: 17); and

5'-TAGAAATGACCTTTCATATGACATTC-3' (primer 4) (SEQ ID NO: 27).

Herein, I represents inosine.

The following primers were used in the second amplification:

5'-GGCCACGCGTCGACTAGTAC-3' (primer 5) (SEQ ID NO: 19); and
5'-TCTGTTTCCATATTGAAAGGCTG-3' (primer 6) (SEQ ID NO: 28). PCR reaction conditions were applied in accordance with protocols attached to the kit.

The amplified 500-bp band was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

(6) 3'-RACE method

The nucleotide sequence on the 3'-side of the DNA fragment obtained by the degenerated PCR was obtained by PCR using a primer prepared based on the information obtained by determination of the nucleotide sequence in (4) above and an oligo dT primer. 3 µl of the first strand cDNA prepared in (2) above was used as a template. The prepared primer was

5'-ATGGTGTCTTATTCAAAGCAAGGCATCGCACA-3' (primer 7) (SEQ ID NO: 29).

Composition of PCR reaction solution:

| | |
|------------------------------|-------|
| Template (first strand cDNA) | 3 µl |
| X10 taq buffer | 5 µl |
| 2.5 mM dNTPs | 4 µl |
| 20 µM primer 7 | 1 µl |
| 10 µM oligo dT primer | 1 µl |
| Milli-Q | 35 µl |
| Taq polymerase (5 U/µl) | 1 µl |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

55°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of 900 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

The obtained full-length nucleotide sequence is shown in SEQ ID NO: 6 of the sequence listing, and the obtained full-length amino acid sequence is shown in SEQ ID NO: 5 thereof. This clone was named MICy.

(7) Expression of protein in *Escherichia coli*

Using a primer produced with a portion corresponding to the N-terminus of the obtained full-length nucleotide sequence of the protein and an oligo dT primer, PCR was carried out employing the first strand cDNA prepared in (2) above as a template.

Primer used:

5'-CGGGATCCGACCATGGTGTCTTATTCAAAGCAAGGCATCGCACA-3' (primer

8) (SEQ ID NO: 30)

Composition of PCR reaction solution:

| | |
|------------------------------|------|
| Template (first strand cDNA) | 3 µl |
| X10 pyrobest buffer | 5 µl |
| 2.5 mM dNTPs | 4 µl |
| 20 µM primer 8 | 1 µl |
| 20 µM oligo dT primer | 1 µl |

Milli-Q 35 μ l

Pyrobest polymerase (5 U/ μ l) 1 μ l

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

55°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of 900 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was subcloned into the BamHI-EcoRI site of a pRSET vector (Invitrogen), and it was then allowed to express in *Escherichia coli* (JM109-DE3). The expressed protein was constructed such that His-tag was attached to the N-terminus thereof, and thus it was purified with Ni-Agarose gel (QIAGEN). Purification was carried out in accordance with the attached protocols. Subsequently, the properties of the purified protein were analyzed.

(8) Analysis of fluorescence properties

Using a solution comprising 20 μ M fluorescent protein (MICy), 150 mM KCl, and 50 mM HEPES(pH 7.4), the absorption spectrum of the protein was measured (Figure 4B). Thereafter, the molar absorption coefficient of the protein was calculated from the value of the peak (472 nm) of the spectrum. The fluorescent protein was diluted with the above buffer solution such that the absorption at 440 nm became 0.001. Its fluorescence spectrum by exciting at 440 nm and its excitation spectrum by a fluorescence at 540 nm were measured (Figure 4A). Likewise, EGFP (CLONTECH) was diluted such that the absorption at 440 nm became 0.001, and its fluorescence spectrum was measured. Setting the quantum yield of EGFP to 0.6, the quantum yield

of the cloned fluorescent protein was obtained. The measurement results are shown in Table 3.

Table 3

| | Excitation maximum | Fluorescence maximum | Molar absorption coefficient | Quantum yield | pH sensitivity | Number of amino acids |
|------|-----------------------|-------------------------|------------------------------------|------------------|-------------------|--------------------------|
| MiCy | 472 nm | 496 nm | 27,250 (472 nm) | 0.90 | pKa = 6.6 | 232 |

(9) Measurement of pH sensitivity

The fluorescent protein was diluted with each of the following buffer solutions. The value of the absorption at 472 nm was determined, and thus the pH sensitivity thereof was measured. The pH levels of the buffer solutions are as follows.

pH 4 and 5: Acetate buffer

pH 6 and 11: Phosphate buffer

pH 7 and 8: HEPES buffer

pH 9 and 10: Glycine buffer

The measurement results are shown in Figure 4C.

(10) Production of pH tolerance mutant MiCy2 of MiCy

Glutamine (Q) at position 166 of MiCy was substituted with histidine (H), so as to produce MiCy2 having stronger fluorescence intensity on the acidic side than that of MiCy (the amino acid sequence of which is shown in SEQ ID NO: 7, and the nucleotide sequence of which is shown in SEQ ID NO: 8). Specifically, pKa of 6.6 was decreased to 5.6, the fluorescence peak was found to be 493 nm, and the excitation peak was found to be 462 nm (Figures 5A and 5B).

Example 4: Isolation of novel fluorescent protein gene (COR) from stony coral, and

analysis of fluorescence properties thereof

(1) Extraction of total RNA

A fluorescent protein gene was isolated from coral. *Montipora* sp. was used as a material. A frozen *Montipora* sp. was crushed in a mortar, and 7.5 ml of “TRIzol” (GIBCO BRL) was then added to 2 g (wet weight) of the crushed *Montipora* sp. Thereafter, the obtained mixture was homogenized and then centrifuged at 1,500 x g for 10 minutes. Thereafter, 1.5 ml of chloroform was added to the obtained supernatant, and the mixture was then stirred for 15 seconds. Thereafter, the mixture was left at rest for 3 minutes. The resultant was then centrifuged at 7,500 x g for 15 minutes. Thereafter, 3.75 ml of isopropanol was added to the obtained supernatant, and the mixture was then stirred for 15 seconds, followed by leaving the mixture at rest for 10 minutes. Thereafter, the resultant was centrifuged at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and 6 ml of 70% ethanol was added to the residue, followed by centrifugation at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and the precipitate was then dissolved in 200 µl of DEPC water. Total RNA dissolved in the DEPC water was 100 times diluted, and the values of O.D.260 and O.D.280 were then measured, so as to determine RNA concentration. As a result, 22 µg of total RNA was obtained.

(2) Synthesis of first strand cDNA

cDNA (33 µl) was synthesized from 4 µg of the total RNA, using a kit for synthesizing first strand cDNA “Ready To Go” (Amersham Pharmacia).

(3) Degenerated PCR

3 µl of the synthesized first strand cDNA (33 µl) was used as a template to carry out PCR.

Primers were produced by making comparison among the amino acid sequences of known fluorescent proteins, extracting similar portions, and converting them to nucleotide sequences.

Primers used:

5'-GAAGGRTGYGTCAAYGGRCAY-3' (primer 1) (SEQ ID NO: 15)

5'-ACVGGDCCATYDGVAAGAAARTT-3' (primer 2) (SEQ ID NO: 16)

I represents inosine; R represents A or G; Y represents C or T; V represents A, C, or G; D represents A, G, or T; S represents C or G; H represents A, T, or C

Composition of PCR reaction solution:

| | |
|-------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 1 | 1 μ l |
| 100 μ M primer 2 | 1 μ l |
| Milli-Q | 35 μ l |
| taq polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 min (PAD)

94°C x 30 sec (denaturation)

52°C x 30 sec (annealing of primers to template)

72°C x 1 min (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 35 times.

72°C x 7 min (final elongation)

4°C (maintenance)

Using 1 μ l of the amplified product obtained in the first PCR reaction as a template, PCR was carried out again under the same above conditions. A 350-bp fragment was cut out via agarose gel electrophoresis, and it was then purified.

(4) Subcloning and sequencing

The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue

white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer. Thereafter, the obtained nucleotide sequence was compared with the nucleotide sequences of other fluorescent protein genes, so as to determine whether the nucleotide sequence of the DNA fragment was derived from a fluorescent protein. With regard to those that were determined to be a part of the fluorescent protein genes, the full-length genes were cloned by the 5'-RACE method and the 3'-RACE method.

(5) 5'-RACE method

In order to determine the nucleotide sequence on the 5'-side of the DNA fragment obtained by the degenerated PCR, the 5'-RACE method was applied using 5'-RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (GIBCO BRL). 5 µg of the total RNA prepared in (1) above was used as a template.

The following primers were used in the first amplification of dC-tailed cDNA: 5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3' (primer 3) (SEQ ID NO: 17); and 5'-CCATCTTCAAAGAGAAAAGACCTTT-3' (primer 4) (SEQ ID NO: 18).

Herein, I represents inosine.

The following primers were used in the second amplification: 5'-GGCCACGCGTCGACTAGTAC-3' (primer 5) (SEQ ID NO: 19); and 5'-CATGAGTTCTTGAAATAGTCAAC-3' (primer 6) (SEQ ID NO: 20). PCR reaction conditions were applied in accordance with protocols attached to the kit.

The amplified 350-bp band was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

(6) 3'-RACE method

The nucleotide sequence on the 3'-side of the DNA fragment obtained by the degenerated PCR was obtained by PCR using a primer prepared based on the information obtained by determination of the nucleotide sequence in (4) above and an oligo dT primer. 3 µl of the first strand cDNA prepared in (2) above was used as a template. The prepared primer was 5'-ATGGCTCTTTCAAAGCACGGTC-3' (primer 7) (SEQ ID NO: 31).

Composition of PCR reaction solution:

| | |
|------------------------------|-------|
| Template (first strand cDNA) | 3 µl |
| X10 taq buffer | 5 µl |
| 2.5 mM dNTPs | 4 µl |
| 20 µM primer 7 | 1 µl |
| 10 µM oligo dT primer | 1 µl |
| Milli-Q | 35 µl |
| Taq polymerase (5 U/µl) | 1 µl |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of approximately 1,000 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified

from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

The obtained full-length nucleotide sequence is shown in SEQ ID NO: 10 of the sequence listing, and the obtained full-length amino acid sequence is shown in SEQ ID NO: 9 thereof. This clone was named COR.

(7) Expression of protein in *Escherichia coli*

Based on the obtained full-length nucleotide sequence, a primer was produced with a portion corresponding to the N-terminus of the protein. An oligo dT primer was used as a primer corresponding to the C-terminal side thereof. Thereafter, using such primers, PCR was carried out employing the first strand cDNA prepared in (2) above as a template.

Primer used:

5'-GGGGGATCCGACCATGGCTCTTTCAAAGCACGGTC-3' (primer 8) (SEQ ID NO: 32)

Composition of PCR reaction solution:

| | |
|------------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 pyrobest buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 8 | 1 μ l |
| 100 μ M oligo dT primer | 1 μ l |
| Milli-Q | 35 μ l |
| Pyrobest polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.
72°C x 7 minutes (final elongation)
4°C (maintenance)

The amplified band with a length of approximately 1,000 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was subcloned into the BamHI-EcoRI site of a pRSET vector (Invitrogen), and it was then allowed to express in *Escherichia coli* (JM109-DE3). The expressed protein was constructed such that His-tag was attached to the N-terminus thereof, and thus it was purified with Ni-Agarose gel (QIAGEN). Purification was carried out in accordance with the attached protocols. Subsequently, the properties of the purified protein were analyzed.

(8) Analysis of fluorescence properties

Using a solution comprising 20 μ M fluorescent protein (COR), 150 mM KCl, and 50 mM HEPES(pH 7.5), the absorption spectrum of the protein was measured (Figure 6B). Thereafter, the molar absorption coefficient of the protein was calculated from the value of the peak (557 nm) of the spectrum. The fluorescent protein was diluted with the above buffer solution such that the absorption at 520 nm became 0.002. Its fluorescence spectrum by exciting at 520 nm and its excitation spectrum by a fluorescence at 600 nm were measured (Figure 6A). Likewise, DsRed2 (CLONTECH) was diluted such that the absorption at 520 nm became 0.002, and its fluorescence spectrum was measured. Setting the quantum yield of DsRed2 to 0.55, the quantum yield of the cloned fluorescent protein was obtained. The measurement results are shown in Table 4.

Table 4

| | Excitation maximum | Fluorescence maximum | Molar absorption coefficient | Quantum yield | pH sensitivity | Number of amino acids |
|-----|-----------------------|-------------------------|------------------------------------|------------------|-------------------|--------------------------|
| COR | 557 nm | 574 nm | 41,750 (557 nm) | 0.41 | pKa < 4.0 | 232 |

(9) Measurement of pH sensitivity

The fluorescent protein was diluted with each of the following buffer solutions to the same concentration. The value of the absorption at 557 nm was determined, and thus the pH sensitivity thereof was measured. The pH levels of the buffer solutions are as follows.

pH 4 and 5: Acetate buffer

pH 6 and 11: Phosphate buffer

pH 7 and 8: HEPES buffer

pH 9 and 10: Glycine buffer

The measurement results are shown in Figure 6C.

Example 5: Isolation of novel chromoprotein gene from sea anemone, and analysis of light-absorbing properties thereof

(1) Extraction of total RNA

A chromoprotein gene was isolated from sea anemone. A single body of *Actinia equina* presenting a red color was used as a material. A frozen *Actinia equina* was crushed in a mortar, and 7.5 ml of "TRIzol" (GIBCO BRL) was then added to 1 g (wet weight) of the crushed *Actinia equina*. Thereafter, the obtained mixture was homogenized and then centrifuged at 1,500 x g for 10 minutes. Thereafter, 1.5 ml of chloroform was added to the obtained supernatant, and the mixture was then stirred for

15 seconds. Thereafter, the mixture was left at rest for 3 minutes. The resultant was then centrifuged at 7,500 x g for 15 minutes. Thereafter, 3.75 ml of isopropanol was added to the obtained supernatant, and the mixture was then stirred for 15 seconds, followed by leaving the mixture at rest for 10 minutes. Thereafter, the resultant was centrifuged at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and 6 ml of 70% ethanol was added to the residue, followed by centrifugation at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and the precipitate was then dissolved in 200 µl of DEPC water. Total RNA dissolved in the DEPC water was 100 times diluted, and the values of O.D.260 and O.D.280 were then measured, so as to determine RNA concentration. As a result, 1.2 mg of total RNA was obtained.

(2) Synthesis of first strand cDNA

cDNA (33 µl) was synthesized from 4 µg of the total RNA, using a kit for synthesizing first strand cDNA “Ready To Go” (Amersham Pharmacia).

(3) Degenerated PCR

3 µl of the synthesized first strand cDNA (33 µl) was used as a template to carry out PCR.

Primers were produced by making comparison among the amino acid sequences of known fluorescent proteins, extracting similar portions, and converting them to nucleotide sequences.

Primers used:

5'- GGI WSB GTI AAY GGV CAY DAN TT -3' (primer 1) (SEQ ID NO: 33); and

5'- GTC ITC TTY TGC ACI ACI GGI CCA TYD GVA GGA AA -3' (primer 2) (SEQ ID NO: 34).

I represents inosine; R represents A or G; Y represents C or T; V represents A, C, or G; D represents A, G, or T; S represents C or G; H represents A, T, or C

Composition of PCR reaction solution:

Template (first strand cDNA) 3 µl

| | |
|-------------------------------|------------|
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 1 | 1 μ l |
| 100 μ M primer 2 | 1 μ l |
| Milli-Q | 35 μ l |
| Taq polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

58°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 35 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

Using 1 μ l of the amplified product obtained in the first PCR reaction as a template, PCR was carried out again under the same above conditions. A 350-bp fragment was cut out via agarose gel electrophoresis, and it was then purified.

(4) Subcloning and sequencing

The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer. Thereafter, the obtained nucleotide sequence was compared with the nucleotide sequences of other fluorescent protein genes, so as to determine whether the nucleotide sequence of the DNA fragment was derived from a fluorescent protein. With regard to those that were determined to be a part of the chromoprotein genes, the full-length genes were cloned by the 5'-RACE method and the 3'-RACE method.

(5) 5'-RACE method

In order to determine the nucleotide sequence on the 5'-side of the DNA fragment obtained by the degenerated PCR, the 5'-RACE method was applied using 5'-RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (GIBCO BRL). 4 µg of the total RNA prepared in (1) above was used as a template.

The following primers were used in the first amplification of dC-tailed cDNA derived from the red color individual:

5'-GGCCACGCGTCGACTAGTACGGGIIIGGGIIIGGGIIG-3' (primer 3) (SEQ ID NO: 17); and

5'-CCT TGA AAA TAA AGC TAT CT-3' (primer 4) (SEQ ID NO: 35).

Herein, I represents inosine.

The following primers were used in the second amplification:

5'-GGCCACGCGTCGACTAGTAC-3' (primer 5) (SEQ ID NO: 19); and

5'-CCC TGT ATG CTT GTG TCC TG-3' (primer 6) (SEQ ID NO: 36). PCR reaction conditions were applied in accordance with protocols attached to the kit.

The amplified 200-bp band was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

(6) Determination of full length nucleotide sequence, and expression of protein in *Escherichia coli*

A primer was produced with a portion corresponding to the N-terminus of the protein obtained in (5) above. An oligo dT primer was used as a primer corresponding to the C-terminal side thereof. Thereafter, using such primers, PCR was carried out employing the first strand cDNA prepared in (2) above as a template.

Primer used:

5'-CCC GGA TCC GAC CAT GGT GTC TTC ATT GGT TAA GAA -3' (primer 7) (SEQ ID NO: 37)

Composition of PCR reaction solution:

| | |
|------------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 pyrobest buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 8 | 1 μ l |
| 100 μ M oligo dT primer | 1 μ l |
| Milli-Q | 35 μ l |
| Pyrobest polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of approximately 900 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was subcloned into the BamHI-EcoRI site of a pRSET vector (Invitrogen), and it was then allowed to express in *Escherichia coli* (JM109-DE3). In addition, the plasmid was recovered, and the full length nucleotide sequence that had been inserted was determined. The clone was named Ume. The full length nucleotide sequence thereof and the full length amino acid sequence thereof are shown in SEQ ID NOS: 12 and 11 of the sequence listing, respectively. The expressed protein was constructed such that His-tag

was attached to the N-terminus thereof, and thus it was purified with Ni-Agarose gel (QIAGEN). Purification was carried out in accordance with the attached protocols. Subsequently, the properties of the purified protein were analyzed.

(7) Analysis of light-absorbing properties

Using a solution comprising 10 μ M chromoprotein (Ume) and 50 mM HEPES (pH 7.9), the absorption spectrum of the protein was measured. Thereafter, the molar absorption coefficient of the protein was calculated from the value of the peak of the spectrum. In the chromoprotein (Ume) derived from the red color individual, the absorption peak was observed at 592 nm (Figure 7A). The measurement results are shown in Table 5.

Table 5

| | Absorption maximum | Molar absorption coefficient | pH sensitivity | Number of amino acids |
|-----|--------------------|------------------------------|------------------------------|-----------------------|
| Ume | 592 nm | 87,000 (592 nm) | Stable between pH 5 and pH10 | 232 |

(9) Measurement of pH sensitivity

The absorption spectrum of the protein was measured in each of the following 50 mM buffer solutions (Figure 7B). The pH levels of the buffer solutions are as follows.

pH 4 and 5: Acetate buffer

pH 6: Phosphate buffer

pH 7 and 8: HEPES buffer

pH 9 and 10: Glycine buffer

The peak value was stable in the range between pH 5 and 10.

Example 6: Isolation of novel fluorescent protein gene from coral (*Lobophytum crassum*)

(1) Extraction of total RNA

A fluorescent protein gene was isolated from coral emitting a fluorescence. *Lobophytum crassum* was used as a material. This coral was crushed with a hammer, and 7.5 ml of "TRIzol" (GIBCO BRL) was then added to 4 g (wet weight) of the crushed *Lobophytum crassum*. Thereafter, the obtained mixture was stirred and then centrifuged at 1,500 x g for 10 minutes. Thereafter, 1.5 ml of chloroform was added to the obtained supernatant, and the mixture was then stirred for 15 seconds. Thereafter, the mixture was left at rest for 3 minutes. The resultant was then centrifuged at 7,500 x g for 15 minutes. Thereafter, 3.75 ml of isopropanol was added to the obtained supernatant, and the mixture was then stirred for 15 seconds, followed by leaving the mixture at rest for 10 minutes. Thereafter, the resultant was centrifuged at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and 6 ml of 70% ethanol was added to the residue, followed by centrifugation at 17,000 x g for 10 minutes. The obtained supernatant was discarded, and the precipitate was then dissolved in 200 µl of DEPC water. Total RNA dissolved in the DEPC water was 100 times diluted, and the values of O.D.260 and O.D.280 were then measured, so as to determine RNA concentration. As a result, 390 µg of total RNA was obtained.

(2) Synthesis of first strand cDNA

cDNA (33 µl) was synthesized from 3 µg of the total RNA, using a kit for synthesizing first strand cDNA "Ready To Go" (Amersham Pharmacia).

(3) Degenerated PCR

3 µl of the synthesized first strand cDNA (33 µl) was used as a template to carry out PCR. Primers were produced by making comparison among the amino acid sequences of known fluorescent proteins, extracting similar portions, and converting them to nucleotide sequences.

Primers used:

5'-GRR AGG IWS BGT HAA YGG VCA-3' (Primer 1) (SEQ ID NO: 38); and

5'-AACTGGAAGAATTCGCGGCCGCAGGAA-3' (Primer 2) (SEQ ID NO :39).

R represents A or G; Y represents C or T; V represents A, C, or G; D represents A, G, or T

Composition of PCR reaction solution:

| | |
|-------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 1 | 1 μ l |
| 100 μ M primer 2 | 1 μ l |
| Milli-Q | 35 μ l |
| Taq polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

72°C x 7 minutes (final elongation)

4°C (maintenance)

Using 1 μ l of the amplified product obtained in the first PCR reaction as a template, PCR was carried out again under the same above temperature conditions. As primers, the following primers were used:

5'-GRR AGG IWS BGT HAA YGG VCA-3' (Primer 1) (SEQ ID NO: 38); and

5'-GTC ITC TTY TGC ACI ACI GGI CCA TYD GVA GGA AA-3' (Primer 3) (SEQ ID NO: 40).

A band with an estimated size of 350 bp was cut out via agarose gel electrophoresis, and it was then purified.

(4) Subcloning and sequencing

The purified DNA fragment was ligated to a pT7-blue vector (Novagen).

Escherichia coli (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of *Escherichia coli*, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer. Thereafter, the obtained nucleotide sequence was compared with the nucleotide sequences of other fluorescent protein genes, so as to determine whether the nucleotide sequence of the DNA fragment was derived from a fluorescent protein. With regard to those that were determined to be a part of the fluorescent protein genes, the full-length genes were cloned by the 5'-RACE method and the 3'-RACE method.

(5) 5'-RACE method

In order to determine the nucleotide sequence on the 5'-side of the DNA fragment obtained by the degenerated PCR, the 5'-RACE method was applied using 5'-RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (GIBCO BRL). 3 µg of the total RNA prepared in (1) above was used as a template.

The following primers were used in the first amplification of dC-tailed cDNA: 5'-GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3' (Primer 4) (SEQ ID NO: 17); and 5'-TTG TCA AGA TAT CGA AAG CGA ACG GCA GAG-3' (Primer 5) (SEQ ID NO: 41).

Herein, I represents inosine.

The following primers were used in the second amplification: 5'-GGCCACGCGTCGACTAGTAC-3' (Primer 6) (SEQ ID NO: 42); and 5'-CTT CTC ACG TTG CAA ATG GC-3' (Primer 7) (SEQ ID NO: 43). PCR reaction conditions were applied in accordance with protocols attached to the kit.

The amplified 600-bp band was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was ligated to a pT7-blue vector (Novagen). *Escherichia coli* (TG1) was transformed therewith, and it was then subjected to blue white selection. Plasmid DNA was purified from white colonies of

Escherichia coli, and the nucleotide sequence of the inserted DNA fragment was determined using a DNA sequencer.

(6) Determination of full length nucleotide sequence, and expression of protein in Escherichia coli

A primer was produced with a portion corresponding to the N-terminus of the protein obtained in (5) above. An oligo dT primer was used as a primer corresponding to the C-terminal side thereof. Thereafter, using such primers, PCR was carried out employing the first strand cDNA prepared in (2) above as a template.

Primer used:

5'-CCC GGA TCC GAT GAG TGT GAT TAC AWC AGA AAT GAA GAT GGA GC
-3' (Primer 8) (SEQ ID NO: 44)

Composition of PCR reaction solution:

| | |
|------------------------------------|------------|
| Template (first strand cDNA) | 3 μ l |
| X10 taq buffer | 5 μ l |
| 2.5 mM dNTPs | 4 μ l |
| 100 μ M primer 8 | 1 μ l |
| 100 μ M oligo dT primer | 1 μ l |
| Milli-Q | 35 μ l |
| Pyrobest polymerase (5 U/ μ l) | 1 μ l |

PCR reaction conditions:

94°C x 1 minute (PAD)

94°C x 30 seconds (denaturation)

52°C x 30 seconds (annealing of primers to template)

72°C x 1 minute (primer elongation)

A cycle consisting of the aforementioned 3 steps was repeated 30 times.

72°C x 7 minutes (final elongation)

4°C (maintenance)

The amplified band with a length of approximately 900 bp was cut out of the gel via agarose gel electrophoresis and then purified. The purified DNA fragment was subcloned into the BamHI-EcoRI site of a pRSET vector (Invitrogen), and it was then allowed to express in *Escherichia coli* (JM109-DE3). In addition, the plasmid was recovered, and the full length nucleotide sequence that had been inserted was determined. The clone was named KnG. The full length nucleotide sequence thereof and the full length amino acid sequence thereof are shown in SEQ ID NOS: 14 and 13 of the sequence listing, respectively.

The expressed protein was constructed such that His-tag was attached to the N-terminus thereof, and thus it was purified with Ni-Agarose gel (QIAGEN). Purification was carried out in accordance with the attached protocols. Subsequently, the properties of the purified protein were analyzed.

(7) Analysis of fluorescence properties

Using a solution comprising 10 μ M fluorescent protein (KnG) and 50 mM HEPES(pH 7.9), the absorption spectrum of the protein was measured (Figure 8A). Thereafter, the molar absorption coefficient of the protein was calculated from the peak value of the spectrum. The absorption peak was observed at 482 nm. The fluorescent protein was diluted with the above buffer solution such that the absorption at 450 nm became 0.005. The fluorescence spectrum obtained by excitation at 450 nm was measured (Figure 8A). Likewise, EGFP (CLONTECH) was diluted such that the absorption at 450 nm became 0.005, and the fluorescence spectrum was measured. Setting the quantum yield of EGFP to 0.6, the quantum yield of the novel protein was obtained. The results are shown in Table 6.

Table 6

| | Excitation maximum | Fluorescence maximum | Molar absorption coefficient | Quantum yield | pH sensitivity | Number of amino acids |
|-----|-----------------------|-------------------------|------------------------------------|------------------|--------------------------------------|--------------------------|
| KnG | 482 nm | 498 nm | 71,000 (482 nm) | 0.41 | Stable between pH4 and pH10 | 224 |

(8) Measurement of pH sensitivity

The fluorescent protein was diluted with each of the following buffer solutions. Thereafter, the fluorescence spectrum was measured. The pH levels of the buffer solutions are as follows.

pH 4 and 5: Acetate buffer

pH 6: MES buffer

pH 7: MOPS buffer

pH 8: HEPES buffer

pH 9 and 10: Glycine buffer

pH 11: Phosphate buffer

The results are shown in Figure 8B.

INDUSTRIAL APPLICABILITY

The present invention provides novel fluorescent proteins derived from *Montipora* sp., *Acropora* sp., and *Lobophytum crassum*. The fluorescent proteins of the present invention are novel proteins having a primary structure different from that of the conventional fluorescent proteins. The fluorescent proteins of the present invention

have certain fluorescence properties, and thus they are useful for molecular biological analysis. That is to say, the use of fluorescent proteins of the present invention enables fluorescent labeling in mammalian cells without exhibiting toxicity. In addition, a mutation is introduced into the fluorescent proteins of the present invention, so as to generate new fluorescence properties.

Moreover, when compared with the wide excitation spectrum of the conventional RFP (DsRed; CLONTECH), the fluorescent protein (COR) of the present invention has a sharper spectrum. Furthermore, a mutation is introduced into the fluorescent protein of the present invention, so as to diversify the fluorescence properties in the red region.

Still further, the present invention provides a novel chromoprotein derived from *Actinia equina*. Since the chromoprotein of the present invention exhibits absorption in the red region and its pH sensitivity thereof is low, it is useful for molecular biological analysis. Still further, since the absorption degree (molar absorption coefficient) of the chromoprotein of the present invention is significantly large, it makes possible to efficiently convert light energy to other types of energy. It is also possible to bring the quantum yield of the chromoprotein of the present invention to close to 1 by genetic modification technology. In such a case, a novel fluorescent protein can be produced.